2

Abstract of the Disclosure

A method uses probabilistic fusion to detect highlights in videos using both audio and visual information. Specifically, the method uses coupled hidden Markov models (CHMMs).

Audio labels are generated using audio classification via Gaussian mixture models (GMMs), and visual labels are generated by quantizing average motion vector magnitudes. Highlights are modeled using discrete-observation CHMMs trained with labeled videos. The CHMMs have better performance than conventional hidden Markov models (HMMs) trained only on audio signals, or only on video frames.